AD_____

Award Number: W81XWH-06-1-0424

TITLE: Characterizing Candidate Oncogenes at 8q21 in Breast Cancer

PRINCIPAL INVESTIGATOR: Jessica Y. Kao

CONTRACTING ORGANIZATION: Stanford University
Stanford, CA 94305

REPORT DATE: March 2008

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
31-03-2008	Annual Summary	1 MAR 2007 - 28 FEB 2008

4. TITLE AND SUBTITLE
Characterizing Candidate Oncogenes at 8q21 in Breast Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-06-1-0424

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jessica Y. Kao

Email: jkao@stanford.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Stanford University
Stanford, CA 94305

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

DNA amplification is a cardinal feature of cancer and plays an important role in tumor progression by altering the gene expression program. These amplified regions are associated with oncogenes of known and unknown identity. By utilizing array comparative genomic hybridization technology to map DNA copy number changes at high resolution, we have identified a recurrent region of high-level amplification at the 8q21 locus that is clinically significant in breast cancer by our preliminary analysis and has yet to be associated with a known oncogene. We have narrowed down this region to two annotated genes with unknown function. In this study we have: (1) created stable cell lines expressing shRNA constructs that will express an siRNA causing inhibition of each candidate oncogene respectively. (2) begun to generate overexpression vectors for each candidate oncogene. (3) raised antibodies to each of the candidate oncogene to further characterize the protein of each candidate oncogenes. Ultimately, this novel oncogene may serve as a new drug target.

15. SUBJECT TERMS
Array CGH

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	12	19b. TELEPHONE NUMBER *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Introduction:

Breast cancer carcinogenesis is caused by molecular genetic changes. These genetic changes ultimately affect the transcriptome. Copy number alterations of the genome is a cardinal feature of cancer and plays an important role in tumor progression by altering the gene expression program. These regions of alteration are associated with oncogenes and tumor suppressor genes of known and unknown identity. Characterization of both CNA's and gene expression profiles have been carried out on breast tumor specimens using microarray technology to gain further insight into the progression of this disease.

We comprehensively characterized both DNA copy number changes and captured the gene expression profiles of 50 breast cancer cell lines, widely used model systems for the study of breast cancer. We found that the cell lines could be classified into three main subtypes, Luminal, Basal A, and Basal B by clustering of gene expression. Overall, this is similar to what is seen in the gene expression analysis of the tumors however, distinct differences were found. Analysis of the copy number profile revealed that the cell lines recapitulated the main genetic changes represented in the tumor data set but contained more changes. Finally, caution should be employed with choosing the appropriate cell line model system when studying specific processes in breast cancer.

Body:

Aim1: Determine if breast cancer subtypes found in tumors are mirrored in commonly used breast cancer cell lines.

To catalog molecular variation in a collection of 50 widely-used breast cancer cell lines, we first profiled gene expression using whole genome oligonucleotide microarrays. Unsupervised hierarchical clustering of the 8,750 most variably expressed genes stratified cell lines into two main groups. One group, designated "luminal" (blue dendrogram branches), contained all the ER-positive cell lines, and was characterized by the expression of ER-alpha regulated genes, as well as genes associated with luminal epithelial differentiation. The other group, designated "basal", contained only ER-negative cell lines and was characterized by the expression of basal epithelial markers including MSN (moesin) and ETS1. Basal cell lines were further stratified into two subgroups, designated A and B.

Our comparisons of expression profiles between breast cancer cell line subtypes and breast tumor subtypes provided valuable information relevant to the suitability of cell lines in modeling known breast tumor heterogeneity. Luminal-A/B tumors best matched luminal cell lines. Notably, basal-like tumors most corresponded to basal-A cell lines. Consistent with this finding, the two breast cancer cell lines from *BRCA1* mutation carriers also clustered in basal-A, where it has been established that *BRCA1*-associated tumors share many features with sporadic basal-like tumors. Interestingly, ERBB2-associated tumors matched both luminal and basal-A lines. While ERBB2 represents a distinct expression subtype in most (but not all) studies, it is noteworthy that most ERBB2 (HER2) cell lines clustered in the luminal subtype. The basis for the discrepant ERBB2 grouping in cell lines and tumors is unclear but warrants further investigation.

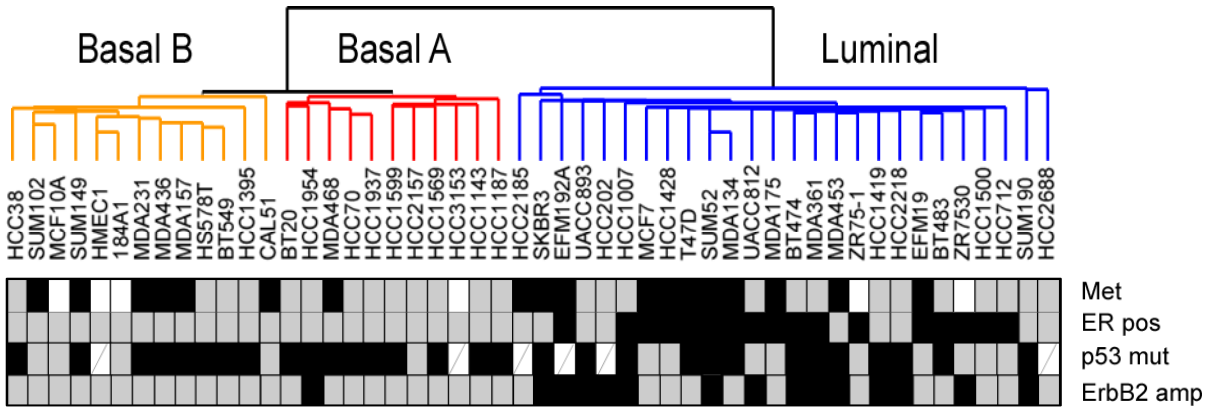

Figure 1. Clustering of expression profiles defines breast cancer cell line subtypes. Clustering stratifies cell lines into two main groups, luminal (blue dendrogram branches) and basal, the latter further subdivided into two subgroups, basal A (red) and basal B (orange). Clinical, pathological and molecular characteristics of cell line expression subtypes. Black boxes indicate metastasis derivation, ER-positivity, *TP53* mutation, *ERBB2*/HER2 positivity. Gray, cross-hatched boxes indicated missing data.

Aim 2: Determine whether breast cancer cell lines accurately depict the genomic changes found in breast cancer tumors.

To survey DNA copy number alterations in the panel of 50 breast cancer cell lines, we carried out CGH on cDNA microarrays with validated performance characteristics and covering 22,000 genes with an average mapping resolution (inter-probe distance) of <70 Kb. Across the sample set, the most frequent CNAs (called by fused lasso method – see Methods) were gains on 1q, 3q, 8q, 17q, and 20q, and losses on 3p, 4, 8p, 9p, and Xq.

Overall, the spectrum of cytoband gains and losses was similar in the cell lines compared to primary tumors, though the frequency of those CNAs was generally higher with the cell lines. Cell line subtype-specific CNAs could be identified by SAM analysis. Luminal cell lines were characterized by more frequent gains on 8q, 12q, 14q, 17q and 20q, and loss on 9p. Of these, gains on 8q and 20q also characterized luminal-B breast tumors, while 17q gain characterized ERBB2-associated tumors. Notably, simple patterns characteristic of luminal-A tumors (1q+, 16p+, 16p-) were not well-represented among the luminal cell lines. Basal-A cell lines were associated by more frequent gains on 3q, 5p, 6p, 9p, 18p, and loss on 14q, of which 16p gain and 14p loss also selectively characterized basal-like breast tumors. Basal-B lines also exhibited characteristic gains/losses, but none shared with basal-like tumors.

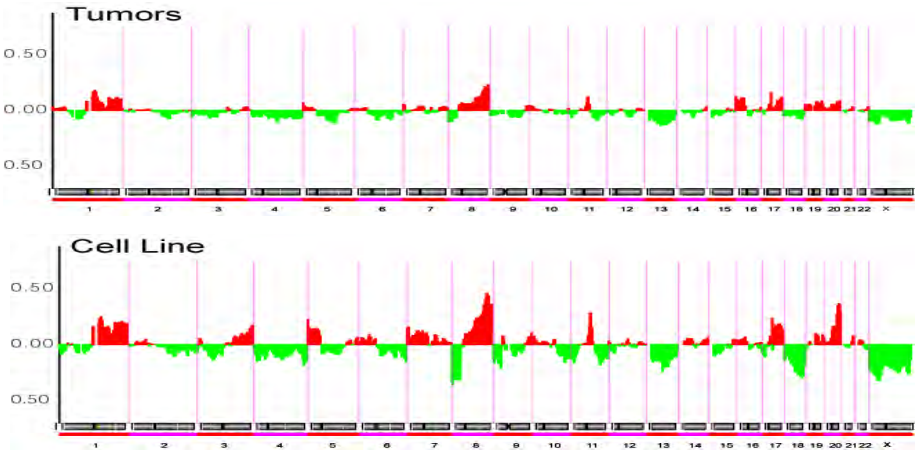

Figure 2. Genomic profiles define spectra of CNAs in cell line subtypes. Spectra of gains (red) and losses (green) across the genome, plotted as average $\log_2$ ratio, for 86 breast tumors, above, compared to the set of 50 cell lines, below.

Aim 3: Determine if copy number specific alterations associated with breast cancer subtypes were also seen in breast cancer cell lines.

Luminal cell lines displayed overall higher frequencies of high-level DNA amplification (i.e. fluorescence ratios ≥ 3, corresponding to at least 5-fold amplification, a characteristic shared with luminal-B tumors. Luminal and basal-A lines both exhibited overall higher frequencies of gain/loss (a characteristic feature of basal-like tumors), compared to basal-B lines. By genomic profiles, luminal cell lines shared features with characteristic of luminal-B tumors, including certain subtype-specific CNAs and overall high levels of DNA amplification. Likewise, basal-A tumors shared features of basal-like tumors, including subtype-specific CNAs and higher levels of gain/loss. However, for both subtypes only a subset of subtype-specific CNAs was preserved. Therefore, at the genomic level it is uncertain how well cell line subtypes faithfully represent tumor subtype counterparts.

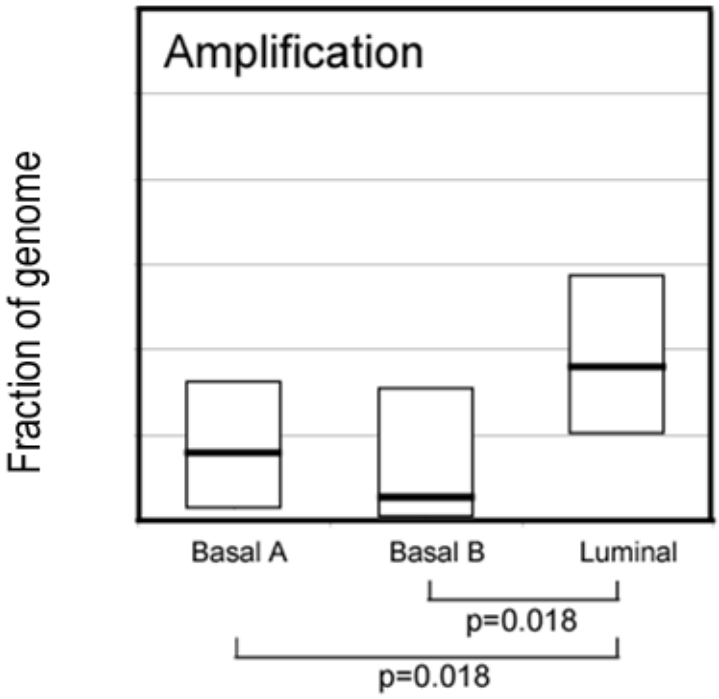

Figure 3. Cell Line subtypes exhibit distinct genomic instabilities
Fraction of genome comprising **(A)** high-level DNA amplification; or **(B)** low-level gain/loss, stratified by cell line subtype (luminal, basal-A, basal-B). Box plots show 25th, 50th and 75th percentiles; *P*-values (Student T-Test) for pairwise comparisons are shown.

Key Research and Training Accomplishments:

- Attended workshop on microarray data analysis software
- Attended American Association of Cancer Research Meeting
- Learned in depth analysis of microarray data
- Completion of several tutorial courses on the Stanford microarray database.
- Teaching Assistant for Immunology and Cancer

Reportable Outcomes:

Kwei K, Kao J, Ratheesh R, Reddy R, Kim Y, Montgomery K, Giacomini C, Choi Y, Chatterjee S, Karikari C, Salari K, Wang P, Hernandez-Boussard T, Swarnalata G, Van de Rijn M, Maitra A, Bashyam M, Pollack J. "Genomic profiling identifies *GATA6* as a lineage-specific oncogene amplified in pancreatobiliary cancer." Plos Genetics (Accepted)

KA Kwei, YH Kim, L Girard, Kao J , M Pacyna-Gengelbach, K Salari, J Lee, Y Choi, M Sato, P Wang, T Hernandez-Boussard, AF Gazdar, I Petersen, JD Minna, JR Pollack. Genomic profiling identifies TITF1 as a lineage-specific oncogene amplified in lung cancer. Oncogene. (2008).

Conclusion:

Transcriptional profiling of breast cancer cell lines identified one luminal and two basal-like (A and B) subtypes. Luminal lines displayed an estrogen receptor (ER) signature and resembled luminal-A/B tumors, basal-A lines were associated with ETS-pathway and BRCA1 signatures and resembled basal-like tumors, and basal-B lines displayed mesenchymal and stem-cell characteristics. Compared to tumors, cell lines exhibited similar patterns of CNA, but an overall higher complexity of CNA (genetically simple luminal-A tumors were not represented), and only partial conservation of subtype-specific CNAs. We identified 51 high-level DNA amplifications and 18 presumptive homozygous deletions, and the resident genes with concomitantly altered gene-expression, highlighting known and novel candidate breast cancer genes. Overall, breast cancer cell lines were genetically more complex than tumors, but retained expression patterns with relevance to the luminal-basal subtype distinction. The compendium of molecular profiles defines cell lines suitable for investigations of subtype-specific pathobiology, biomarkers and therapies, and provides a resource for discovery of new breast cancer genes.

Appendix:

Curriculum Vitae

Jessica Kao

650-776-1180 • jkao@stanford.edu • 240 Flynn Ave., Mountain View, CA 94043

Education

Stanford University ***2000 – March 2008***
PhD in Cancer Biology
Thesis: Identification of novel oncogenes in breast cancer by aCGH and microarray gene expression analysis.
Cumulative 3.7 GPA

Emory University ***1996 - 2000***
B.S. Summa Cum Laude in Biology
B.A. Music
Cumulative 3.7 GPA

Relevant Coursework:
Oral Communication for Graduate Students
Teaching Oral Communication

Research Experience

Stanford University ***2003 - 2007***
PhD candidate, laboratory of Jonathan Pollack
Project: Identification of novel oncogenes in breast cancer by aCGH and microarray gene expression analysis.

- Performed aCGH and gene expression profiling on 50 beast cancer cell lines.
- Performed analysis on parallel DNA/RNA data set.
- Assisted in standardizing new protocol for gene expression on Stanford oligo arrays.
- Consulted with and taught new lab members aCGH protocol.
- Troubleshot standardized, and taught fellow lab members assays for cell proliferation, cell cycle analysis by FACS, siRNA gene knockdown, and apoptosis.
- Extensive experience with validation of array results by qRT-PCR.

Stanford University ***2000 - 2003***
PhD candidate, laboratory of Nic Denko
Project: Characterization of novel hypoxia induced genes

Emory University ***1997 - 2000***
Undergraduate Research, laboratory of Anita Corbett
Project: Cell cycle analysis in *Saccharomyces cerevisiae*

Emory University ***1996 - 1997***
Undergraduate Research, laboratory of Grey Crouse
Project: DNA mismatch repair in *Saccharomyces cerevisiae*

Center for Disease Control – Atlanta, Ga. ***1996 - 1997***
Research Assistant
Project: Diagnostic testing by PCR

Walt Disney World Cancer Research Institute – Orlando, Fl ***1995 - 1996***
Research Assistant

Project: Characterizing CML cell line, K-562

Teaching Experience

Oral Communications Consultant ***2005 – Present***

Center for Teaching and Learning, Stanford University

Initiated, designed, and delivered new workshops on oral communication and presentations at the undergraduate and graduate level. Advised and coached undergraduate and graduate students on presentation, interviewing, and speaking skills. Trained other tutors on effective use of powerpoint and teaching oral communication to scientists. Workshops include: Command an audience through posture and body alignment; Common Presentation Pitfalls.

Teaching Assistant

Stanford University

Genomics 109Q: A Technical and Cultural Revolution ***Winter 2007***
Cbio 101 (Cancer Biology) ***Spring 2006***
Bio 42 (Genetics, Biochemistry, and Molecular Biology) & ***Winter 2000***
Led weekly undergraduate discussion sections. Designed, created, and graded assignments.

Emory University

Bio 143 (Genetics) ***Spring 1998***
Led weekly undergraduate discussion sections. Held office hours to aid in students learning.

Dance Instructor

Dance 4 Health non profit ***2006 – Present***

Designed curriculum for after school programs to improve the health and fitness of at risk individuals for obesity and childhood diabetes by integrating nutritional facts with ballroom dance inspired creative movement. Currently developing teacher training curriculum.

Stanford University ***2001 - 2005***

Designed and taught social and competitive ballroom dance classes including: waltz, tango, and salsa.

Leadership Experience

Team Captain, Dance Competition Organizer ***2000 - 2006***

Stanford Ballroom Dance Team, Stanford University

Created and organized the implementation of a successful competitive ballroom dance program. Coordinated and organized the planning and execution of the annual ballroom dance competition. Increased revenue by 15% for 3 years in a row. Doubled the attendance from 500 to 1000 spectators the final year. Documented and trained future organizers. Advised other western US collegiate teams in organizational process.

Executive Board Member, Publicity Chair ***2005 - 2006***

USA Dance– Northern California Chapter

Assisted in the organization and planning of regional ballroom dance competitions and the 2006 National Ballroom Dancesport Championships. Managed ticket box office, handled customer service issues, and trained volunteers over the 3-day national competition. Increased exposure and attendance of the event through newspaper articles and free advertising. Responsible for media relations.

Freshman Advisor ***Fall 2005***

Stanford Farm Mentor, Stanford University

Advised students on both academic and personal issues pertaining to college life.

Board Member ***2001 - 2002***

Graduate Housing Advisory Committee, Stanford University

Liaison between graduate students and the university. Gathered information about student's housing needs through surveys. Presented conclusions to the university committee.

Publications

Kao, J. and Pollack J.R. (2006) "RNA interference-based functional dissection of the 17q12 amplicon in breast cancer reveals contribution of co-amplified genes." Genes, Chromosome, and Cancer; 45(8):761-9.

Abstracts and Presentations

Kao, J. and Pollack J.R. (2007) "The Functional Dissection of Amplicons in Breast Cancer by RNAi." 3rd Annual Asia Pacific Multidisciplinary Metting for Cancer Genomics Research, Hong Kong.

Kao, J. and Pollack J.R. (2005) "Interrogating localized regions of high amplification in breast cancer." Stanford Cancer Biology Conference, Asilomor, Ca.

Kao, J. and Corbett, A. (2000), Mutational analysis of the *Saccharomyces cerevisiae* cyclin-dependent kinase, CDC28. Symposium conducted at The Southeastern Regional Yeast Meeting Birmingham, AL.

Kao, J. and Corbett, A. (1999), Identification of a novel gene implicated in cell cycle regulation in *Saccharomyces cerevisiae*. Symposium conducted at the 39th Annual American Society of Cell Biology Meeting, Washington DC.

Awards and Honors

Department of Defense Breast Cancer Research Pre-Doctoral Training Grant.	***2005***
NIH Cancer Biology Training Grant	***2000-2005***
Howard Hughes Pre-doctoral Fellowship Honorable Mention	***2000***
ODK Leadership Honor Society	***2000***
Mortar Board Senior Honor Society	***2000***
Phi Sigma Biology Honor Society	***2000***
Mu Phi Music Honor Society	***2000***
Howard Hughes Summer Undergraduate Research Fellow	***1998***

Skills

Lab: aCGH on Stanford cDNA and oligo arrays • microarray gene expression analysis • knockdown of gene expression by siRNA • quantitative RT-PCR • FACS analysis • FISH • IHC • immunofluorescence microscopy • tissue culture • western blotting • northern blotting • molecular cloning • standard molecular biology techniques

Computer: Microsoft Excel • Powerpoint • Word • Adobe Photoshop • Illustrator